DESIGN OF FRAMEWORK FOR INFORMATION RETRIEVAL FROM ONLINE SOCIAL NETWORK

SHORT SYNOPSIS

Submitted in partial fulfillment of the requirement of the degree of

DOCTOR OF PHILOSOPHY

to

Manav Rachna International Institute of Research and Studies
(Deemed to be University)

by

MOHIT CHOWDHARY
Registration number:17/PhD/053

Under the supervision of
Dr. POONAM CHAHAL
Associate Professor,
Faculty of Engineering and Technology

FACULTY OF ENGINEERING AND TECHNPLOGY

MANAV RACHNA INTERNATIONAL INSTITUTE OF RESEARCH & STUDIES

SECTOR-43, SURAJKUND –DELHI ROAD, FARIDABAD – 121001

FACULTY OF ENGINEERING AND TECHNOLOGY

Certificate

I hereby declare that the short synopsis "Design of Framework for Information Retrieval from Online Social Network" being submitted in partial fulfillment of the requirement for the degree of Doctor of Philosophy in Computer Science and Engineering under faculty of Engineering and Technology of Manav Rachna International Institute of Research & Studies, Faridabad during the academic year 2019-20.

It is a bona fide record of original work carried out under the supervision of Dr Poonam Chahal and has not been submitted in full or part to any other university or institute for degree or diploma. Further, it is certified that short synopsis is plagiarism free and proof has been submitted.

(MOHIT CHOWDHARY)
17/PhD/053

Signature of the Supervisor
Dr Poonam Chahal, Associate Professor
Department of Computer Sc & Engineering
Faculty of Engineering and Technology

Signature of the Joint-Supervisor
Full name, designation and faculty

(Chairman, DRC)

Chairperson, CDP

ABSTRACT

In today's era, the use of internet is increasing at a rapid rate. The network of networks i.e. Internet in which a lot of information is available for the access of various users on a click. As the time passes technology is also growing, as a result, World Wide Web been switching to Web 2.0. The Web 2.0 is a network of social websites in which people are connected to each other for sharing information to enhance their skills, online purchase, electronic items and food items etc. The social network helps the users to express the views related to current affairs, current trends in fashion, different food habits, feedback of services availed, goods purchased, politics etc. This leads to generation of huge amount of information which may be structured, semi-structured and unstructured.

A number of techniques have been given by different researchers to find semantic similarity and page ranking for query expansion. Some researchers have used various social activities and time series data of users for retrieval of information from various social web sites. Many researchers have considered features for semantic information retrieval like metadata, social book marking, Tagmap matrix, social proximity etc. but still there exist gaps that need to be overcome for efficient retrieval of information. Some of the major gaps that needs to be worked on like the social annotations embedded in the social network data and the extraction of relevant features. The other Major gap in the consideration of relationship between the concepts and enhancing the same using temporal analysis.

This research work will be carried out to design a framework which will process the data present on social network for the efficient retrieval of semantic information. For processing of social information the Natural Language Processing (NLP) techniques will be applied for detecting the concepts present in social information along with the relationship between them. The detection of interrelated concepts will be done by using a structured framework like ontologies which will further help in retrieving the semantic information as per the expectations of user.

Key words: World Wide Web, Web Search, Information Retrieval, Natural Language Processing, Online Social Network.

CONTENTS

S. No	Description	Page No
1	Introduction 1.1 World Wide Web 1.2 Web Standards 1.3 Social Websites 1.3.1 Blogs and Wikis 1.3.1 Mobile Device Enabled Social Sites 1.3.1 Content based sites with social features 1.3.1 Interest Based Websites	1-6
2	Literature Review	7-8
3	Description of Broad Area 3.1 Description 3.1.1 Information retrieval 3.1.2 Online Social networks 3.1.3 Social Web Search 3.1.4 Social Information retrieval 3.2 Research Gaps 3.3 Research Problem	9-14
4	Objectives of the Study	15
5	Methodology to be adopted	16
6	Expected Outcome of the research	17
7	References	18-20

1. Introduction

The World Wide Web (WWW) is basically the resource of information in which all the users are accessing the available information by using the Hypertext Transfer Protocol (HTTP). There is an international community which is acting as World Wide Web Consortium (W3C) including various members working on the development of the standards of Web. It may be defined as a: "The universe of interconnected networks which are used for accessing the information over this connected network of systems, and an embodiment of human knowledge."These interlinked web pages contains huge amount of information in the form of web pages that are accessed by using the tool i.e. search engine. Search engine primarily contains three components Indexer, Ranker and Crawler. Website contains a very large amount of information in web pages that are accessed using Internet. This helps the users of web to gather and display desired information available on web in the form of text.

1.1 World Wide Web

The web is incorporated with the major ideas that focus on serving the users of the web for the retrieval of information, sharing information widely, create forums for technical enhancement, E commerce, Banking, etc. With the use of advancement in technology the WWW has become Dynamic, Interactive and Evolving day by day. The web browser is the basic tool that is used to provide the interface to many users for the access of information available on the web. The use of open standards like TCP/IP, HTML, CSS etc and the open source has become the need of the era. The information present on the WWW is in the form of web pages written in natural language that is difficult to be processed by the machine. The information is linked with each other by using the hypertext and Cross-Platform and it is widely distributed over nearly 200 million sites which are active as of December 2018. The basic architecture of WWW consists of the following:

Hypertext: It is the part of text, provides a way to jump from one part of current document to another part of document through internet connections which is provided in the text itself. The hypertext mentioned present in the web documents is popularly known as hyperlinks.

Resource Identifiers: These are the URL/URI which identifies a resource on a server or internet.

Client-Server Computing: A setup of interconnected systems which is used to perform computing or communication among these systems, as one system at a time acts as a client and other system act as a server. Client side computer makes requests to server side computer which will provide resources or services to client side computer.

Markup language: The set of codes which are assigned to elements of a text and which need to be embedded in text is done by using Markup language. These texts indicates structure, semantic, or advice available on a document.

1.2 Web Standards

In general, the internet or web consists of three major standards which are been discussed as follows:

Uniform Resource Identifier (URI), An identifier used to identify and locate a resources present on the Web, for accessing information from Website pages;

HyperText Transfer Protocol (HTTP), a set of rules which denote the communication between the browser and websites.

HyperText Markup Language (HTML), is a markup language used to represent information present in hypertext documents in structured manner.

1.3 Social Websites

A Social web is known as a set of relations present on websites having social behavior by which people are connected to each other by means of website. It circumscribes working of websites and software running on these social websites that are used in order to maintain especially interconnections and to incorporate the communication between the interconnected individuals. These types of online social connections and interactions constitute the basis of the online activities such as shopping online, learning, gaming, professional, social media etc. The users who are having similar taste and interests web 2.0 has been a facilitator for the interaction between people.

The users or peoples who are connected to the social media are constantly posting the reviews and providing the forum to the society for sharing views on large platform through the social web or social media platforms. The influence of social network is very large and it is ever changing over the course of period. By the increase in the activities of people on the Web and

due to increase in communication, the information regarding their social behavior become more viable and accessible to the external world or in public.

In today's world, billions of users are available on the internet that are accessing social networking websites to remain connected with friends, associates, employers, may even discover new "friends", to share their views. Most of the common sites based on social platforms such as Linkedin, Twitter, MySpace, Facebook, future Dataweb, facilitate people working in diverse organizations and those who are connected to each other. The user generates content in different forms like photographs, videos, blogs, etc. To enhance the accessibility of the social media, a lot of users are also using the mobile platforms by making use of various application provided by different platforms.

In 2018, by the report of Facebook there are 2400 million members, accessing their account on Facebook. as compared to 2008. The availability of users on MySpace was recorded as 100 million and on the other hand YouTube had videos and channels as 100 million and 2.9 million respectively. The mentioned no shows a great expansion of online social web users till 2018 and it's still enhancing in exponential manner. In general, social websites are rapidly enhancing and updating, by moving beyond traditional websites that are used to connect individuals and give an entirely new experience to the user's, which lead to new way of living a life with numerous social connections.

Social sites like Classmates along with SixDegrees, are probly first social sites that were founded in 1995 and 1997 respectively. The first social site for writing an online diary was started in 1998 and named as open diary. Till now Open Diary has hosted nearly millions of diaries in digital formats. This community was competent enough to bring jointly the writers of online Diary to a common platform. It was one of the most common social networking communities which leads to the emergence of a new term like weblog and later on this term evolved as blog.

Open Diary has played an enormous role in the evolution of today's social websites, which are now popularly known as social media or social networking. With the increase in speed of internet and availability to the masses, social networking has gained popularity which leads to the numerous social networking sites, some of the major common websites like Myspace, facebook, Twitter and Instagram etc. With the evolution of social sites, the web has become a platform for facilitating many users to exchange and share info world wide web.

1.3.1 Blogs and Wikis

New age collaborations are emerging in today's internet era like blogs and Wikis , having features like interacting in groups and allow to communicate simultaneously to each other. Blogs are acting as a discussion forum in between the different users having same interest and allowing the user to share their views, Wikis allow any contributor to edit it and who has been granted access to the content created by the contributor of Blogs and Wikis can see the recent discussions and add their views over these and they are allowed to make changes in the content, blogs and Wikis can be accessed by the non members. Primary purpose of using Blog and Wiki services is to make use of information available with them and share to educate all the members and non members of social web platform. The social interactions generally focused on the content of topic and views associated with the same. The information provided by Blogs and Wikipedia and other Wiki are used by both types of user, those who are contributing in the content addition and those people who reference them as resources of information for their use and processing. Blogs follow a mechanism in which they allow their members to express their ideas and thoughts of their own interest on current and other topics and allow other members to comment on those ideas; on the other hand Wikis facilitate group collaboration related with same interest and thoughts. These tools i.e. blogs and wikis broadly open a gateway for the users to share their thoughts and ideas. Most frequent users of these sites are academicians, research scholar etc working in same vicinity who are having same requirement to achieve the goal of enhancing learning, research, innovative skills etc. This also helps different users of social media to get different interpretation of same topics and subjects. They share the resources which are not approachable or available to through other modes.

1.3.2 Mobile Device Enabled Social Sites

By increasing interest and usability of social web services, online web services are providing interconnectivity round the clock. Now days so social media sites are now offering support on other available platforms such as mobile devices and internet phone. Twitter, Facebook, YouTube and other popular websites are coming up with new style of content sharing and posting to account in which users are allowed to post content from other sites and sources directly to their social media platform. This includes all the activities such as updating in their

status, receiving updates from their friends etc. This can easily be done via mobile devices or any other devices having internet access. The social sites also providing convenience to the customers through mobile devices are generally termed as mobile device enabled social sites. These mobile enabled social sites help customers increase their contacts and friends online. Social web applications available on mobile platform needs more enhanced technologies like augmented reality related with games that will provide users wider experience which was earlier implemented in SCVGR and LAYAR.

1.3.3 Content based sites with Social features

Web sites that are primarily built for specific purposes such as Online Shopping, hosting of video content, online music apps etc. are generally specified as content based web sites. These content based sites add features that allow discussions about the quality of content, feedback for the content available or request for some other product, services etc. This results in establishing collaboration to expand the account of users to a maximum level for the benefit of organization growth. The major online retailers such as Amazon had implemented content based features, to collect the reviews of their valuable customer, in order to enhance their sales, customer base and ultimately great success in business.

The users of amazon who wish to purchase any product seek Customer reviews that help customers to know more about a product by getting reviews of other persons. These reviews provided by the customers, provide a lot of valuable information, which help the users for getting idea or experience about the product. These customer reviews are more like a magnet for an organization, which are used as a tool to expand the customers for business growth.

1.3.4 Interest Based Web Sites

Interest websites have come up with social features and they are implementing this social web features to their sites to broaden their appeal in their own customers and future targeted customers.

Allrecipes.com, community of people, of all age groups, who likes cooking and share their ideas to prepare recipes in different styles, with the users of this website. Along with this they exchange recipes with one another, other users can rate their recipes, can post reviews about recipes which are being tried by them, provide suggestions about the improvements in and

alter them accordingly. The reviews, ratings and recommendations are a valuable resource for this community, as it provides feedback from the website users and their families about the recipes. These websites are connected with other networking websites, Twitter and Instagram and blogs, to dilate its presence over the social web.

2. Literature Review

Lioma, C., Blanco, R., and Moens M.F. (2009) have discussed non personalized Query Expansion for retrieval of relevant documents with the help of Social tags, freely available on social tagging system. The authors formulated Query Expansion as a logical reference. The mechanism to identify key terms were added and integrated into the query with the help of approximation of uncertainty to find relevance of original query terms.

Lin, Y., Lin, H., Jin, S. and Ye, Z. (2011) have given relevant feedback data collected by using social annotations. The social annotation analyzed to get the key terms for construction of process of query expansion. These key terms with the help of machine learning method helps to carry out personalized Query Expansion.

Biancalana, C., Micarelli, A. and Squarcella, C. (2008) have developed a novel system named Nereau. this system was developed to carry a personalized query Expansion, selection of most suitable terms based on, metadata. These relevant terms retrieved from Social bookmarking to overcome lack of Semantic Attribute, were used to increase Co-Occurrence of terms in a document.

Bertier, M., Guerraoui, R., Leroy, V. and Kermarrec, A.M. (2009) have given new algorithm named as TagRank algorithm. This algorithm determines the relevancy of tag in Query Expansion. The authors used PageRank algorithms and TagMap matrix, in which distance is computed by considering different tags. These tags are first mapped with keys and respective items.

Biancalana, C., Gasparetti, F., Micarelli, A. and Sansonetti, G. (2013) have discussed a method of Social and personalized query expansion with the help of Three Dimensional Matrix. In 3D one more dimension was added by the authors from traditional approach in which 2D co-occurrence Matrix is used. The third dimension was added by using semantic classes having all the semantic property and folksonomies, which are extracted from social bookmarking.

Bouadjenek, M.R., Hacid, H. and Bouzeghoub, M. (2016) used methodology of ranking for expension of query the methodology is being used by computing the similarity in context of semantic. In addition, the query composition, tags and the proximity among the specified query is also considered.

Bouadjenek, M.R., Hacid, H. and Bouzeghoub, M. (2019) focused on social activities of the user for calculating the score for individual document. The computed score of document is known as Personalized Social Document Representation (PSDR). This score is considered for the available data set for the retrieval of information as per user needs.

Zhang, Z. and Lan, M. (2014) have given ranking model built by using supervised machine learning techniques to perform re-ranking. The relevant tweets are retrieved with relevant information by extracting semantic features from the feedback of pseudo relevance and results of Google search engine.

Karpagam, P. and Sivasubramanian, S. (2016) focused on new concepts predicted accurately by applying temporal content in the social media text messages. This is done by using semantic computation in each series of time.

Rao, J., Yang, W., Zhang, Y., Ture, F. and Lin, J. (2018) have given neural ranking model for the ranking of short messages on social media. It makes use of hierarchical convolution layers in the learning of characters, words and phrases in order to find the latent semantic soft match relevance signals for them. It integrates the evidence collected from various sources that include matches between query, URLs contained in the post.

Zhu, X., Huang, J., Zhou, B., Li, A. and Jia, Y. (2017) designed a framework for twitter using real time data, the authors also used a Boolean logic for filtering keywords to improve the ranking performance of tweets. It works on the enhancement of the accuracy to get superior quality tweets. In order to explore user preferences properly, a query expansion is done by using external search engine. It integrates semantic features and social attributes of tweets and also adopts dynamic strategy to provide only useful tweets as desired by the users.

Vigneshwari, S. and Aramudhan, M. (2015) suggested a model used for multiple ontologies. It is used to retrieve useful information by building taxonomy within the concepts and explore the relationships between them.

3. Description of Broad Area

3.1 Description

Social websites are rich source of information as user are accessing social sites or websites equipped with social features on their mobile devices. The content which is available on social sites is of vast use i.e Recommender System, Temporal data retrieval, and feedback systems can be designed by analyzing information available on these sites.

3.1.1 Information Retrieval

It is a process of finding pertinent information from documents available in the form of structured, unstructured and semi structured (commonly as text) which are used to satisfy the requirement of gathering information from a large collection which is usually stored on computers and other digital storage media.

The process of collecting pertinent information which is gathered and displayed as per the requirements of users having different kinds of users e.g. paralegals, librarians, and professional searchers of similar interest. With the change in technology and requirements, millions of users and likewise professionals are busy in the process of information reclamation. When users perform a search through search engines or even in their email they are involved in the process of information retrieval knowingly or unknowingly. The domain of information retrieval is growing at a very fast face, and now a days it is most promising form of information gathering and accessing over common database style searching. The process of Information Retrieval consist of of data and information. They are namely Structured, Unstructured and Semi structured. The data which does not have clear, semantically evident, and it is also not easy for a system to process is Unstructured data.

The process of Extracting the relevant information from a set of documents with a user query is performed, in order to find the relevant documents. The main characteristics of information retrieval process.

Accuracy: The data which is used for the process of extracting relevant information is of large size, and it contain many types of shortcomings in accordance with the amount of relevant information extracted from these documents. Having a good accuracy in information extracted

from these documents is not possible which mean missing a lot of relevant texts or content in text.

Speed: Due to large quantity of input text available and process that text, the speed of processing, the data or information can lead to a reverse bottleneck. Practically, prior to performing an information retrieval process, data has to be processed using either NLP techniques or run as "batch" indexing tool to minimize the computation time.

Consistency: The accuracy of data is decreased for the reason of information retrieval techniques which intern require indexing of the text by the authors itself or by a group of indexers to retrieve the information. The probable inconsistencies I text could be removed by carrying out an automatic processing to the data.

Ease of use: with the increase in the availability of personal computer systems, and trained intermediary, the earlier model information retrieval are obsolete now a days, which result into the design of systems which are having a high priority responsivenes.

3.1.2 Online Social Network

Online social network is considered as social plateform which consists of various network build by social arrangement which constitutes academicians, business personnel and other categories of persons which different purposes. The users of OSN (online social network) connects with each other for different types of communication to get update on recent trends or activities of their interest.

The objects, users and relations on online social network are managed for maintaiing the overall structure of social network. The table 1.1 shows the examples of social network also with the short description of different social sites.

Table 1.1: Social Networking Sites With Introduction

Year (Start)	Social Website	Short Introduction
1997	SixDegrees.com	Used for public profile and friend list
1999	LiveJournal.com	Blog host
2000	Xanga.com	Having features of Clone of Live Journal
2000	SkyBlog.com	Clone of Live Journal
2001	CyWorld.com	Social networking site
2001	Ryze.com	Social network for business purposes
2002	Friendster.com	Social network site
2003	LinkedIn.com	For Professional linkages
2003	Tribe.net	Main focus on classifieds but later on adopted by the Burning Man community
2004	MySpace.com	A competitor of friendster
2004	Orkut.com	Social network by Google
2004	Mixi.jp	Social networking service for Japan
2004	Lunarstorm.se	Social networking service for Sweden
2004	Grono.net	Social networking service for Poland
2004	Hi5.com	Dating website
2004	Xing.com (openBC)	business oriented social network for Germany/Europe
2004	Bebo.com	Social Networking
2004	flicker.com	An image hosting service and video hosting service
2004	CiteUlike.com	Based on Social Bookmarking
2005	Facebook.com	Social network for students
2005	YouTube.com	Video sharing website
2005	studivz.net	Social network site for students
2006	Twitter	Micro blogging site with social network features
2010	Instagram.com	Photo and video-sharing social networking service

The OSN structure is basically having contest of relations which are explicitly or explicitly formed by the OSN users.

To analyze and understand the structure and working procedures of the OSN there are so many techniques and methods commonly known as Social network analysis. Generally a social network have three major relationships like some people follow a no of other people who are having same interest or common objectives. These relations are

i) Follower and Following

ii) ii) Publisher and Subscriber

iii) iii) Friend and Friendship.

So there are three types of categories of relationships

i) Symmetric Relations

ii) ii) Asymmetric Relations

iii) iii) K-Partite Relationships

3.1.3 Social Web Search

Users are generating content such as news, reviews, comments, images and videos on social platform, such as, Instagram, Flicker, Facebook, Twitter and LinkedIn, which are related with the queries that are to be searched over traditional search engines. This mechanism or behavior of user to retrieve, search information on social search engine is known as social web search. This mechanism used already available algorithms with enhanced algorithms that are designed by various researchers to be implemented on unstructured data.

In general social network contain various forms of social relationships. These relations on OSN are formed by various groups of professionals, individual, academicians, researchers, students, business personnel etc. In addition to this personnel groups for official communication are also formed by various peoples for informal communications

3.1.4 Social Information Retrieval

Due to the advancement in technology and structure of websites, lead to increase in number of users and the applications which are available on the web are now more interactive to their sources. This leads to generation in an enormous amount of information on the social web. The procedure of retrieving meaningful information available on social network is consist of

relationships among the different people and content shared by them. This leads to derive and describe the model for social relationship. The activity of IR is performed due to the concern of users to satisfy their information needs. Social Information Retrieval process primarily targets to provide relevant information and related content in the specific domains. The recommendation, information retrieval, and research, which is covering broad topics such as social network analysis, collaborative filtering, Q&A systems, and subjective relevance judgments.

3.2 Research Gaps

To perform information retrieval from online social network, there are a number of methods which has been given by various researchers. Some uses semantic similarity and page ranking for query expansion. Other researchers considered social activities, time series data etc for retrieving information from social web. Many other features like social proximity, metadata, TagMap matrix, social bookmarking is also considered for relevant semantic information retrieval in social web. But still there are gaps which need to overcome for effective content extraction from OSN.

Some of the gaps are as follows:

1. Social Annotations done for query expansion generally uses personalized and non personalized approach this limits the coverage of key terms, behavior, attributes for relevant feature extraction which can be overcome by using machine learning approaches.
2. Semantic Similarity has been considered by the various researchers for extracting semantic information from social web. But the researchers considered majorly the concepts rather than relationship in concepts which are to be considered for the results as per user expectations. This can be done efficiently by using ontology. This is basically conceptualization of view.
3. The consideration of conceptual Relationships using temporal analysis is required for predicting the temporal content in social media text messages.

3.3 Research Problem

To design a framework (Model) to extract semantic and meaningful information available on OSN using Natural Language Processing (NLP) techniques. The efficient techniques will be applied for the retrieval of concepts and relationship between concepts using structured framework like ontologies. The interrelated concepts with features be used to retrieve the semantic information.

4. Objectives of the Study

1. To collect the desired information from various sources of OSN.
2. To apply Natural Language Processing (NLP) techniques for processing of collected information from OSN.
3. To design an extended model for OSN using SIR models.
4. To apply enhanced model on OSN to extract the user desired information.
5. To contrast the designed model with existing SIR model.

5. Methodology to be adopted

1. An extensive Literature Survey will be performed to get in-depth sight of the IR are.
2. An analysis for ranking of query, extraction of content, indexing of content etc of numerous IR and SIR model will be done.
3. NLP techniques will be implemented for processing of information on OSN.
4. For efficient extraction of content and information from OSN Further Ontology will be embedded.
5. An enhanced model using the conceptualization of concepts i.e. Ontology will be applied.

6. Expected Outcome of the Research

After getting the processed data, an enhanced model will be designed for application on OSN for efficient content retrieval by applying natural language processing technique.

The model will process the supplied data and retrieve the relevant semantic information contained on social sites.

7. References

Bertier, M., Guerraoui, R., Leroy, V. and Kermarrec, A.M., 2009, March. Toward personalized query expansion. In *Proceedings of the Second ACM EuroSys Workshop on Social Network Systems* (pp. 7-12).

Biancalana, C., Gasparetti, F., Micarelli, A. and Sansonetti, G., 2013. Social semantic query expansion. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *4*(4), p.60.

Biancalana, C., Micarelli, A. and Squarcella, C., 2008, October. Nereau: a social approach to query expansion. In *Proceedings of the 10th ACM workshop on Web information and data management* (pp. 95-102).

Bouadjenek, M.R., Hacid, H. and Bouzeghoub, M., 2016. Social networks and information retrieval, how are they converging? A survey, a taxonomy and an analysis of social information retrieval approaches and platforms. *Information Systems*, *56*, pp.1-18.

Bouadjenek, M.R., Hacid, H. and Bouzeghoub, M., 2019. Personalized Social Query Expansion Using Social Annotations. In *Transactions on Large-Scale Data-and Knowledge-Centered Systems XL* (pp. 1-25).

Bouadjenek, M.R., Hacid, H., Bouzeghoub, M. and Daigremont, J., 2011, July. Personalized social query expansion using social bookmarking systems. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval* (pp. 1113-1114).

Bouadjenek, M.R., Hacid, H., Bouzeghoub, M. and Vakali, A., 2013, July. Using social annotations to enhance document representation for personalized search. In *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval* (pp. 1049-1052).

Brusilovsky, P., Smyth, B. and Shapira, B., 2018. Social search. In *Social Information Access* (pp. 213-276).

Efron, M., Lin, J., He, J. and De Vries, A., 2014, July. Temporal feedback for tweet search with non-parametric density estimation. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval* (pp. 33-42).

Ferguson, P., O'Hare, N., Lanagan, J., Phelan, O. and McCarthy, K., 2012, April. An investigation of term weighting approaches for microblog retrieval. In *European Conference on Information Retrieval* (pp. 552-555). Springer, Berlin, Heidelberg.

Halpin, H., Tuffield, M., Appelquist, D., Brickley, D., Carvahlo, M., Iannella, R., Passant, A., Perey, C. and Story, H., 2014. A standards-based, open and privacy-aware social web. W3C Incubator Group Report.

Karpagam, P. and Sivasubramanian, S., 2016. Identifying the Impact of Semantic Similarity in Semantic Social Media through Evolving Term for Effective Web Mining. *Indian Journal of Science and Technology, 9*(44).

Lau, C.H., Li, Y. and Tjondronegoro, D., 2011, November. Microblog Retrieval Using Topical Features and Query Expansion. In *TREC*.

Lin, Y., Lin, H., Jin, S. and Ye, Z., 2011, July. Social annotation in query expansion: a machine learning approach. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval* (pp. 405-414).

Lioma, C., Blanco, R. and Moens, M.F., 2009, September. A logical inference approach to query expansion with social tags. In *Conference on the Theory of Information Retrieval*(pp. 358-361). Springer, Berlin, Heidelberg.

Massoudi, K., Tsagkias, M., De Rijke, M. and Weerkamp, W., 2011, April. Incorporating query expansion and quality indicators in searching microblog posts. In *European Conference on Information Retrieval* (pp. 362-367). Springer, Berlin, Heidelberg.

Rao, J., Yang, W., Zhang, Y., Ture, F. and Lin, J., 2018. Multi-perspective relevance matching with hierarchical convnets for social media search. *arXiv preprint arXiv:1805.08159*, pp.7-11.

Vigneshwari, S. and Aramudhan, M., 2015. Social information retrieval based on semantic annotation and hashing upon the multiple ontologies. *Indian Journal of Science and Technology*, *8*(2), p.103.

Ward, R., 2006. Blogs and wikis: A personal journey. *Business Information Review*, *23*(4), pp.235-240.

Wasserman, S. and Faust, K., 1994. *Social network analysis: Methods and applications* (Vol. 8). Cambridge university press.

Zhang, Z. and Lan, M., 2014. Estimating Semantic Similarity between Expanded Query and Tweet Content for Microblog Retrieval. In *TREC*.

Zhu, X., Huang, J., Zhou, B., Li, A. and Jia, Y., 2017. Real-time personalized twitter search based on semantic expansion and quality model. *Neurocomputing*, *254*, pp.13-21.

SHORT SYNOPSIS MOHIT CHOWDHARY Sep20,2019

ORIGINALITY REPORT

7%	0%	1%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

Excluding matches less than 14 words

PRIMARY SOURCES

1	Submitted to Manav Rachna International University	5%
2	Claudio Biancalana, Fabio Gasparetti, Alessandro Micarelli, Giuseppe Sansonetti. "Social semantic query expansion", ACM Transactions on Intelligent Systems and Technology, 2013 Publication	1%
3	Submitted to University of St. Gallen Student Paper	1%
4	Mohamed Reda Bouadjenek, Karin Verspoor. "Multi-field query expansion is effective for biomedical dataset retrieval", Database, 2017 Publication	<1%
5	Submitted to Farmington High School Student Paper	<1%
6	Submitted to Bridgepoint Education Student Paper	<1%

Vikram M. Bhatia
University Librarian
Dr. O.P. Bhatia Central Library
Manav Rachna International
Institute of Research & Studies
Faridabad

Exclude quotes	On	Exclude matches	< 14 words
Exclude bibliography	On		